

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/670,568C
Source: 1FW16
Date Processed by STIC: 12/7/04

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 12/07/2004

PATENT APPLICATION: US/09/670,568C

TIME: 08:09:42

Input Set : A:\09-670,568 Sequence Listing.txt

Output Set: N:\CRF4\12072004\I670568C.raw

3 <110> APPLICANT: IKAWA, Yoji
 4 IKAWA, Shuntaro
 5 OBINATA, Masuo
 7 <120> TITLE OF INVENTION: HUMAN P51 GENES AND GENE PRODUCTS THEREOF
 9 <130> FILE REFERENCE: Q61014
 11 <140> CURRENT APPLICATION NUMBER: 09/670,568C
 C--> 12 <141> CURRENT FILING DATE: 2001-01-18
 14 <150> PRIOR APPLICATION NUMBER: JP 10-100467
 15 <151> PRIOR FILING DATE: 1998-03-27
 17 <150> PRIOR APPLICATION NUMBER: PCT/JP99/01512
 18 <151> PRIOR FILING DATE: 1999-03-24
 20 <160> NUMBER OF SEQ ID NOS: 29
 22 <170> SOFTWARE: PatentIn version 3.2
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 448
 26 <212> TYPE: PRT
 27 <213> ORGANISM: Homo sapiens
 30 <220> FEATURE:
 31 <221> NAME/KEY: DOMAIN
 32 <222> LOCATION: (1)..(59)
 33 <223> OTHER INFORMATION: transactivation domain
 35 <220> FEATURE:
 36 <221> NAME/KEY: DNA_BIND
 37 <222> LOCATION: (142)..(321)
 38 <223> OTHER INFORMATION: DNA binding domain
 40 <220> FEATURE:
 41 <221> NAME/KEY: DOMAIN
 42 <222> LOCATION: (353)..(397)
 43 <223> OTHER INFORMATION: oligomerization domain
 45 <400> SEQUENCE: 1
 47 Met Ser Gln Ser Thr Gln Thr Asn Glu Phe Leu Ser Pro Glu Val Phe
 48 1 5 10 15
 51 Gln His Ile Trp Asp Phe Leu Glu Gln Pro Ile Cys Ser Val Gln Pro
 52 20 25 30
 55 Ile Asp Leu Asn Phe Val Asp Glu Pro Ser Glu Asp Gly Ala Thr Asn
 56 35 40 45
 59 Lys Ile Glu Ile Ser Met Asp Cys Ile Arg Met Gln Asp Ser Asp Leu
 60 50 55 60
 63 Ser Asp Pro Met Trp Pro Gln Tyr Thr Asn Leu Gly Leu Leu Asn Ser
 64 65 70 75 80
 67 Met Asp Gln Gln Ile Gln Asn Gly Ser Ser Thr Ser Pro Tyr Asn
 68 85 90 95
 71 Thr Asp His Ala Gln Asn Ser Val Thr Ala Pro Ser Pro Tyr Ala Gln

RAW SEQUENCE LISTING

DATE: 12/07/2004

PATENT APPLICATION: US/09/670,568C

TIME: 08:09:42

Input Set : A:\09-670,568 Sequence Listing.txt

Output Set: N:\CRF4\12072004\I670568C.raw

```

72          100          105          110
75 Pro Ser Ser Thr Phe Asp Ala Leu Ser Pro Ser Pro Ala Ile Pro Ser
76          115          120          125
79 Asn Thr Asp Tyr Pro Gly Pro His Ser Phe Asp Val Ser Phe Gln Gln
80          130          135          140
83 Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr Ser Thr Glu Leu Lys
84 145          150          155          160
87 Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro Ile Gln Ile Lys Val
88          165          170          175
91 Met Thr Pro Pro Pro Gln Gly Ala Val Ile Arg Ala Met Pro Val Tyr
92          180          185          190
95 Lys Lys Ala Glu His Val Thr Glu Val Val Lys Arg Cys Pro Asn His
96          195          200          205
99 Glu Leu Ser Arg Glu Phe Asn Glu Gly Gln Ile Ala Pro Pro Ser His
100          210          215          220
103 Leu Ile Arg Val Glu Gly Asn Ser His Ala Gln Tyr Val Glu Asp Pro
104 225          230          235          240
107 Ile Thr Gly Arg Gln Ser Val Leu Val Pro Tyr Glu Pro Pro Gln Val
108          245          250          255
111 Gly Thr Glu Phe Thr Thr Val Leu Tyr Asn Phe Met Cys Asn Ser Ser
112          260          265          270
115 Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu Ile Ile Val Thr Leu
116          275          280          285
119 Glu Thr Arg Asp Gly Gln Val Leu Gly Arg Arg Cys Phe Glu Ala Arg
120          290          295          300
123 Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala Asp Glu Asp Ser Ile
124 305          310          315          320
127 Arg Lys Gln Gln Val Ser Asp Ser Thr Lys Asn Gly Asp Gly Thr Lys
128          325          330          335
131 Arg Pro Phe Arg Gln Asn Thr His Gly Ile Gln Met Thr Ser Ile Lys
132          340          345          350
135 Lys Arg Arg Ser Pro Asp Asp Glu Leu Leu Tyr Leu Pro Val Arg Gly
136          355          360          365
139 Arg Glu Thr Tyr Glu Met Leu Leu Lys Ile Lys Glu Ser Leu Glu Leu
140          370          375          380
143 Met Gln Tyr Leu Pro Gln His Thr Ile Glu Thr Tyr Arg Gln Gln Gln
144 385          390          395          400
147 Gln Gln Gln His Gln His Leu Leu Gln Lys His Leu Leu Ser Ala Cys
148          405          410          415
151 Phe Arg Asn Glu Leu Val Glu Pro Arg Arg Glu Thr Pro Lys Gln Ser
152          420          425          430
155 Asp Val Phe Phe Arg His Ser Lys Pro Pro Asn Arg Ser Val Tyr Pro
156          435          440          445
159 <210> SEQ ID NO: 2
160 <211> LENGTH: 2816
161 <212> TYPE: DNA
162 <213> ORGANISM: Homo sapiens
165 <220> FEATURE:
166 <221> NAME/KEY: exon

```

RAW SEQUENCE LISTING

DATE: 12/07/2004

PATENT APPLICATION: US/09/670,568C

TIME: 08:09:42

Input Set : A:\09-670,568 Sequence Listing.txt

Output Set: N:\CRF4\12072004\I670568C.raw

```

167 <222> LOCATION: (145)..(1488)
169 <220> FEATURE:
170 <221> NAME/KEY: polyA_signal
171 <222> LOCATION: (2786)..(2791)
173 <400> SEQUENCE: 2
174 tcgttgatat caaagacagt tgaaggaaat gaattttgaa acttcacggt gtgccaccct 60
176 acagtactgc cctgaccctt acatccagcg ttctgtagaa acccagctca tttctcttgg 120
178 aaagaaagtt attaccgatc cacc atg tcc cag agc aca cag aca aat gaa 171
179 Met Ser Gln Ser Thr Gln Thr Asn Glu
180 1 5
182 ttc ctc agt cca gag gtt ttc cag cat atc tgg gat ttt ctg gaa cag 219
183 Phe Leu Ser Pro Glu Val Phe Gln His Ile Trp Asp Phe Leu Glu Gln
184 10 15 20 25
186 cct ata tgt tca gtt cag ccc att gac ttg aac ttt gtg gat gaa cca 267
187 Pro Ile Cys Ser Val Gln Pro Ile Asp Leu Asn Phe Val Asp Glu Pro
188 30 35 40
190 tca gaa gat ggt gcg aca aac aag att gag att agc atg gac tgt atc 315
191 Ser Glu Asp Gly Ala Thr Asn Lys Ile Glu Ile Ser Met Asp Cys Ile
192 45 50 55
194 cgc atg cag gac tcg gac ctg agt gac ccc atg tgg cca cag tac acg 363
195 Arg Met Gln Asp Ser Asp Leu Ser Asp Pro Met Trp Pro Gln Tyr Thr
196 60 65 70
198 aac ctg ggg ctc ctg aac agc atg gac cag cag att cag aac ggc tcc 411
199 Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Gln Asn Gly Ser
200 75 80 85
202 tcg tcc acc agt ccc tat aac aca gac cac gcg cag aac agc gtc acg 459
203 Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser Val Thr
204 90 95 100 105
206 gcg ccc tcg ccc tac gca cag ccc agc tcc acc ttc gat gct ctc tct 507
207 Ala Pro Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Ala Leu Ser
208 110 115 120
210 cca tca ccc gcc atc ccc tcc aac acc gac tac cca ggc ccg cac agt 555
211 Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro His Ser
212 125 130 135
214 ttc gac gtg tcc ttc cag cag tcg agc acc gcc aag tcg gcc acc tgg 603
215 Phe Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala Thr Trp
216 140 145 150
218 acg tat tcc act gaa ctg aag aaa ctc tac tgc caa att gca aag aca 651
219 Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala Lys Thr
220 155 160 165
222 tgc ccc atc cag atc aag gtg atg acc cca cct cct cag gga gct gtt 699
223 Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Pro Gln Gly Ala Val
224 170 175 180 185
226 atc cgc gcc atg cct gtc tac aaa aaa gct gag cac gtc acg gag gtg 747
227 Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr Glu Val
228 190 195 200
230 gtg aag cgg tgc ccc aac cat gag ctg agc cgt gaa ttc aac gag gga 795
231 Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn Glu Gly
232 205 210 215

```

RAW SEQUENCE LISTING

DATE: 12/07/2004

PATENT APPLICATION: US/09/670,568C

TIME: 08:09:42

Input Set : A:\09-670,568 Sequence Listing.txt

Output Set: N:\CRF4\12072004\I670568C.raw

234	cag att gcc cct cct agt cat ttg att cga gta gag ggg aac agc cat	843
235	Gln Ile Ala Pro Pro Ser His Leu Ile Arg Val Glu Gly Asn Ser His	
236	220 225 230	
238	gcc cag tat gta gaa gat ccc atc aca gga aga cag agt gtg ctg gta	891
239	Ala Gln Tyr Val Glu Asp Pro Ile Thr Gly Arg Gln Ser Val Leu Val	
240	235 240 245	
242	cct tat gag cca ccc cag gtt ggc act gaa ttc acg aca gtc ttg tac	939
243	Pro Tyr Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Val Leu Tyr	
244	250 255 260 265	
246	aat ttc atg tgt aac agc agt tgt gtt gga ggg atg aac cgc cgt cca	987
247	Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg Arg Pro	
248	270 275 280	
250	att tta atc att gtt act ctg gaa acc aga gat ggg caa gtc ctg ggc	1035
251	Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val Leu Gly	
252	285 290 295	
254	cga cgc tgc ttt gag gcc cgg atc tgt gct tgc cca gga aga gac agg	1083
255	Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg Asp Arg	
256	300 305 310	
258	aag gcg gat gaa gat agc atc aga aag cag caa gtt tcg gac agt aca	1131
259	Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp Ser Thr	
260	315 320 325	
262	aag aac ggt gat ggt acg aag cgc cgg ttt cgt cag aac aca cat ggt	1179
263	Lys Asn Gly Asp Gly Thr Lys Arg Pro Phe Arg Gln Asn Thr His Gly	
264	330 335 340 345	
266	atc cag atg aca tcc atc aag aaa cga aga tcc cca gat gat gaa ctg	1227
267	Ile Gln Met Thr Ser Ile Lys Lys Arg Arg Ser Pro Asp Asp Glu Leu	
268	350 355 360	
270	tta tac tta cca gtg agg ggc cgt gag act tat gaa atg ctg ttg aag	1275
271	Leu Tyr Leu Pro Val Arg Gly Arg Glu Thr Tyr Glu Met Leu Leu Lys	
272	365 370 375	
274	atc aaa gag tcc ctg gaa ctc atg cag tac ctt cct cag cac aca att	1323
275	Ile Lys Glu Ser Leu Glu Leu Met Gln Tyr Leu Pro Gln His Thr Ile	
276	380 385 390	
278	gaa acg tac agg caa cag caa cag cag cag cac cag cac tta ctt cag	1371
279	Glu Thr Tyr Arg Gln Gln Gln Gln Gln Gln His Gln His Leu Leu Gln	
280	395 400 405	
282	aaa cat ctc ctt tca gcc tgc ttc agg aat gag ctt gtg gag ccc cgg	1419
283	Lys His Leu Leu Ser Ala Cys Phe Arg Asn Glu Leu Val Glu Pro Arg	
284	410 415 420 425	
286	aga gaa act cca aaa caa tct gac gtc ttc ttt aga cat tcc aag ccc	1467
287	Arg Glu Thr Pro Lys Gln Ser Asp Val Phe Phe Arg His Ser Lys Pro	
288	430 435 440	
290	cca aac cga tca gtg tac cca tagagcccta tctctatatt ttaagtgtgt	1518
291	Pro Asn Arg Ser Val Tyr Pro	
292	445	
294	gtgttgatt tccatgtgta tatgtgagtg tgtgtgtgtg tatgtgtgtg cgtgtgtatc	1578
296	tagccctcat aaacaggact tgaagacact ttggctcaga gacccaactg ctcaaaggca	1638
298	caaagccact agtgagagaa tcttttgaag ggactcaaac ctttacaaga aaggatgttt	1698
300	tctgcagatt ttgtatcctt agaccggcca ttggtgggtg aggaaccact gtgtttgtct	1758

RAW SEQUENCE LISTING

DATE: 12/07/2004

PATENT APPLICATION: US/09/670,568C

TIME: 08:09:42

Input Set : A:\09-670,568 Sequence Listing.txt

Output Set: N:\CRF4\12072004\I670568C.raw

```

302 gtgagctttc tgttgtttcc tgggagggag gggtcaggtg gggaaagggg cattaagatg 1818
304 tttattggaa cccttttctg tcttcttctg ttgttttctt aaaattcaca gggaagcttt 1878
306 tgagcaggtc tcaaacttaa gatgtctttt taagaaaagg agaaaaaagt tgttattgtc 1938
308 tgtgcataag taagttgtag gtgactgaga gactcagtca gaccctttta atgctggtca 1998
310 tgtaataata ttgcaagtag taagaaacga aggtgtcaag tgtactgctg ggcagcgagg 2058
312 tgatcattac caaaagtaat caactttgtg ggtggagagt tctttgtgag aacttgcatt 2118
314 atttgtgtcc tccctcatg tgtaggtaga acattttctta atgctgtgta cctgcctctg 2178
316 ccactgtatg ttggcatctg ttatgctaaa gtttttcttg tacatgaaac cctggaagac 2238
318 ctactacaaa aaaactgttg tttggccccc atagcaggtg aactcatttt gtgcttttaa 2298
320 tagaaagaca aatccacccc agtaatattg cccttacgta gttgtttacc attattcaaa 2358
322 gctcaaaaata gaatttgaag ccctctcaca aaatctgtga ttaatttgct taattagagc 2418
324 ttctatccct caagcctacc taccataaaa ccagccatat tactgatact gttcagtgca 2478
326 tttagccagg agacttacgt tttgagtaag tgagatccaa gcagacgtgt taaaatcagc 2538
328 actcctggac tggaaattaa agattgaaag ggtagactac ttttcttttt ttactcaaa 2598
330 agtttagaga atctctgttt ctttccattt taaaaacata ttttaagata atagcataaa 2658
332 gactttaaaa atgttctctc cctccatctt cccacaccca gtcaccagca ctgtattttc 2718
334 tgtcaccaag acaatgattt cttgttattg aggctgttgc ttttgtggat gtgtgatttt 2778
336 aattttcaat aaacttttgc atcttggttt aaaagaaa 2816

```

339 <210> SEQ ID NO: 3

340 <211> LENGTH: 393

341 <212> TYPE: PRT

342 <213> ORGANISM: Homo sapien

344 <400> SEQUENCE: 3

```

346 Met Glu Glu Pro Gln Ser Asp Pro Ser Val Glu Pro Pro Leu Ser Gln
347 1 5 10 15
350 Glu Thr Phe Ser Asp Leu Trp Lys Leu Leu Pro Glu Asn Asn Val Leu
351 20 25 30
354 Ser Pro Leu Pro Ser Gln Ala Met Asp Asp Leu Met Leu Ser Pro Asp
355 35 40 45
358 Asp Ile Glu Gln Trp Phe Thr Glu Asp Pro Gly Pro Asp Glu Ala Pro
359 50 55 60
362 Arg Met Pro Glu Ala Ala Pro Arg Val Ala Pro Ala Ala Pro
363 65 70 75 80
366 Thr Pro Ala Ala Pro Ala Pro Ala Pro Ser Trp Pro Leu Ser Ser Ser
367 85 90 95
370 Val Pro Ser Gln Lys Thr Tyr Gln Gly Ser Tyr Gly Phe Arg Leu Gly
371 100 105 110
374 Phe Leu His Ser Gly Thr Ala Lys Ser Val Thr Cys Thr Tyr Ser Pro
375 115 120 125
378 Ala Leu Asn Lys Met Phe Cys Gln Leu Ala Lys Thr Cys Pro Val Gln
379 130 135 140
382 Leu Trp Val Asp Ser Thr Pro Pro Pro Gly Thr Arg Val Arg Ala Met
383 145 150 155 160
386 Ala Ile Tyr Lys Gln Ser Gln His Met Thr Glu Val Val Arg Arg Cys
387 165 170 175
390 Pro His His Glu Arg Cys Ser Asp Ser Asp Gly Leu Ala Pro Pro Gln
391 180 185 190
394 His Leu Ile Arg Val Glu Gly Asn Leu Arg Val Glu Tyr Leu Asp Asp
395 195 200 205

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/670,568C

DATE: 12/07/2004

TIME: 08:09:43

Input Set : A:\09-670,568 Sequence Listing.txt

Output Set: N:\CRF4\12072004\I670568C.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date